AFRL-IF-RS-TR-2005-102
Final Technical Report
March 2005

INFRASTRUCTURE VULNERABILITY ASSESSMENT AND DEFENSE

University of Massachusetts at Amherst

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2005-102 has been reviewed and is approved for publication

APPROVED: /s/

KEVIN A. KWIAT
Project Engineer

FOR THE DIRECTOR: /s/

WARREN H. DEBANY, JR., Technical Advisor
Information Grid Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	MARCH 2005	Final Dec 02 – Nov 04

4. TITLE AND SUBTITLE
INFRASTRUCTURE VULNERABILITY ASSESSMENT AND DEFENSE

5. FUNDING NUMBERS
C - F30602-03-2-0008
PE - 63789F
PR - AIPT
TA - GA
WU - 01

6. AUTHOR(S)
Lixin Gao

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Massachusetts at Amherst
Department of Electrical and Computer Engineering
Knowles Engineering Building
Amherst Massachusetts 01003

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/IFGA
525 Brooks Road
Rome New York 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2005-102

11. SUPPLEMENTARY NOTES
AFRL Project Engineer: Kevin A. Kwiat/IFGA/(315) 330-1692/ Kevin.Kwiat@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*
The goal of this Infrastructure Vulnerability Assessment and Defense Project was to develop models for characterizing the worse-case scenarios for the Internet's routing infrastructure. In addition, sub-goals were to propose counter-measures to these vulnerabilities, implement, and experiment the proposed counter-measures, and evaluate their potential impact.

14. SUBJECT TERMS
Internet, Packet Loss, Routing, IP, Path Switching

15. NUMBER OF PAGES
27

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents

List of Figures

List of Tables

1. Introduction

This project was a "seedling" effort to generate new research directions in understanding the inherited vulnerability of protocols, implementations and topologies in the current Internet. As the "network of networks", the Internet is owned and controlled by different companies or organizations. The connectivity and communications within and among networks are "glued" together by a set of protocols defined by the Internet research communities, which are modified and improved over decades of time. Those protocols are usually designed in the principle to achieve simplicity; allow maximum flexibilities for the network operators and essentially a free hand for the equipment manufactures to implement them. Although the Internet has been successfully operating for more than 20 years and continues attracting more and more popularity, the inherited vulnerability by the distributed nature of the network has never been fully understood.

The Distributed Deny of Service (DDoS) attack is one of the many inherited security vulnerabilities that are unique in an environment like the Internet. Attackers may remotely control hundreds of computers to launch attacks on a single machine simultaneously. The resources on the victim computer will be quickly depleted and unable to perform its normal task. The propagation of Internet worms is another such example. The worm program duplicates itself and propagates to other computers in the Internet usually by exploiting implementation loophole of some network protocols. The speed of the worm propagation is so fast that it generates traffic that may disrupt normal network operation or even bring down backbone routers and links. In both cases, the security vulnerabilities are easy and cheap to be exploited by the attacker, while hard to prevent or defend by network operators or researchers. Such attacks will be impossible or too expensive to implement on a conventional network such as PSTN.

The existence of those vulnerabilities poses a bigger question of how the basic principles on which the Internet was built on may create security vulnerabilities that are hard to avoid and may be exploited by malicious attackers. Are there vulnerabilities that are hidden in normal operation cases but may appear or exacerbate in those abnormal scenarios such as a malicious attack? If we can identify such vulnerabilities, can we estimate the potential destruction or disruption caused by those vulnerabilities?

It is important to understand those problems, as we are increasingly dependent on the Internet for service, business and a source of information. We hope our research may help to improve the reliability and the end-to-end performance experienced by the users; and increase security and reliability for network service in other scenarios such as abnormal network traffic or malicious attacks. In addition, such research may shed light on the future design of better and more secure Internet protocols.

As an initial "seedling" effort, our goal was to examine a variety of potential problems of Internet, which may cause operational problems or security vulnerabilities. Of the approaches we explored, the two detailed in this report are the directions we matured the most during the duration of this project. The first direction is targeted at end-to-end performance such as packet losses. We make the effort to understand its cause, impact

on user applications and propose mechanisms to improve end-to-end loss performances. We start with investigating the characteristics of the end-to-end packet loss based on the active measurements we performed on a wide-area Internet testbed. We find that besides network congestion, Internet routing instability such as the BGP or IGP changes may be the cause of packet losses. In particular, the bursty losses seen by the end users account for a significant portion of the total packet loss, and exhibit higher correlations with routing instability. On top of these observations, we propose several algorithms that perform end-to-end path switching based on routing instability information. Our results show that path switching performed by the hosts at the end of an Internet path may effectively reduce the overall end-to-end loss rate and significantly reduce the loss burstness.

The second direction we described here is that we initiated the investigation on understanding the routing loops in the Internet. Similar to the approaches in our first effort, we use active end-to-end measurements to observe the prevalence and the characteristics of routing loops on the Internet. We are especially interested in the routing loops that persist over a long period of time and those routing loops that extend across multiple networks, because those routing loops have a more significant impact on network operations.

Efforts in the above directions have yielded results that indicate promising further developments. By understanding the characteristics of routing instability and routing abnormalities such as the routing loops, we may be able to propose mechanisms to avoid performance degradation and increase network availability when networks undergo normal routing fluctuation or malicious attacks. Eend-to-end path switching may be seen as the initial step towards such a goal.

1.1 Publications

Papers describing portions of the work in this project were submitted to ACM SIGMETRICS 2005.

1.2 Patents

There were no patents filed as a result of this project yet.

1.3 Financial Summary

The full budget of $70,000 was expended during the performance period. This money was used primarily to support graduate students working on the project, as well as faculty members involved in this project.

2. Packet Loss on Internet Paths

2.1 Objective

In this study, we set out to investigate whether Internet routing instability is one of the reasons for impairing network performance - such as packet loss. Previous studies have tried to model the loss patterns in the Internet and find ways to improve network resilience and performance [1 2 3]. It is shown that predicting Internet packet loss by monitoring loss patterns in real-time is possible [4]. The success of such predictions is based on the underlying fact that packet loss in the Internet exhibits a degree of temporal correlation and therefore is predictable from simply a statistical point of view. However, predicting Internet packet loss based on past loss observations does not reflect the underlying cause of those packet losses. To this end, we study the characteristics and correlation between the Internet routing instability and packet loss using network measurement techniques. Based on the result, we propose new loss avoidance mechanism to improve the end-to-end performance and network resilience.

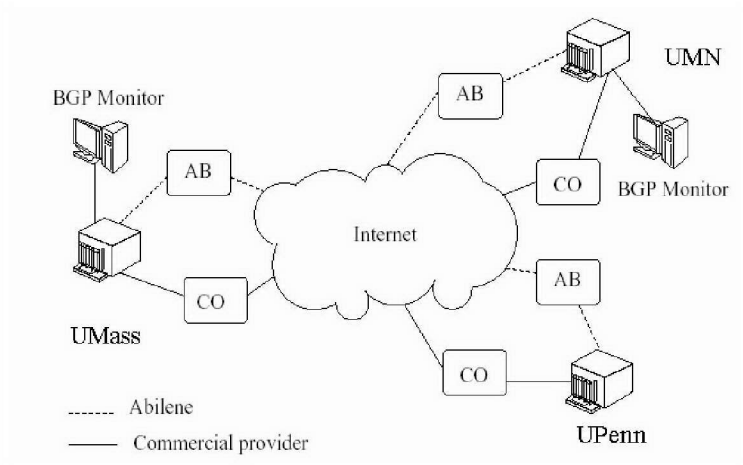


Figure 1. Wide-area network testbed architecture

2.2 Testbed and Experiment Setup

We performed our basic packet loss observation and route change pattern on a wide-area network using an active measurement approach.

The architecture of our wide area testbed is shown in Figure 1. The testbed consists of nodes located at three major universities in the U.S. Two of the three universities are located on the east coast (University of Massachusetts and University of Pennsylvania),

and one in the mid-west (University of Minnesota). Each of the three sites is a multi-homed network and can select the Abilene network (AB) as well as one other commercial provider (CO) to reach the global Internet.

At each site, a workstation is configured as a gateway node. The gateway nodes were assigned multiple IP addresses, so that it may choose which provider network to deliver outgoing traffic by working together with the border routers. At all three universities, the border routers were configured with special routing policies to enable automatic route selection. The border routers at UPenn and UMN select the outgoing provider according to the source address of the IP packets, while UMass border routers install two static routing entries to select outgoing providers according to the destination addresses of the IP packets. IP tunnels are created between the testbed nodes so that the IP packets sent to the tunnel will automatically be encapsulated with proper source and destination addresses. Our measurement programs are only required to select the correct IP tunnel in order to choose which provider to forward outgoing traffic.

We also collected BGP routing information at UMass and UMN using separate workstations running Zebra BGP daemons. Those BGP monitors peer with the edge router of the campus network and regularly record and archive the updates received from their providers. UMN uses BGP to connect to its both providers, while UMass connects to Abilene using BGP and uses static default route to connect to its commercial provider (Verio). The clocks on all the gateway nodes and the BGP monitors are synchronized using the Network Time Protocol (NTP), except for the gateway node at UPenn.

We performed continuous UDP probes on all the provider paths and monitor end-to-end packet loss information. We also performed traceroute on all the paths among testbed nodes to detect IP level path change. The UDP probes are sent every 1-second, and the intervals between consecutive traceroute probes are 10 seconds. Our analysis is based on the measurement data we collected during a 10-day long period.

2.3 Related Work

There have been several previous studies on the loss characteristics in the Internet. In [1 2], Yanjnik et al. investigated the temporal correlation of packet losses, especially in a multicast network setting. The existence of such temporal correlation indicates that it is possible to predict Internet packet loss based on past loss observation. In [5], Feamster et al. pointed out the correlation between BGP instability and network failure, based on the measurement data collected from an Internet testbed of 31 topologically diversified hosts. The result of that study suggests that passive monitoring BGP messages can predict about 20% of impending failures. Therefore reactive routing systems may use that information to provide some degree of resilience to Internet path failures. In [4], Tao et al. studied the end-to-end delay and loss performance on parallel paths between node pairs in the Internet. The result shows that there usually is a dominant path that outperforms the other parallel paths between a pair of Internet nodes in terms of end-to-end delays. This is because the propagation, instead of the queuing delay is usually the major contribution of the overall end-to-end delay. On the other hand, end-to-end loss shows similar

characteristics on parallel Internet paths, regardless of the geographic locations of the end nodes.

In our analysis, we adopt a measurement approach in much finer granularity to capture the routing dynamics than that is performed at [5]. Unlike [5], which only focuses on major network failures on Internet paths; our work focuses on the characteristics of end-to-end packet loss. On the other hand, Tao's work [4] emphasizes the temporal statistics of losses on Internet paths, while our study focuses on the relationships between routing dynamics and end-to-end loss, with particular emphasis on bursty losses.

2.4 Loss observation

Table 1 shows the average loss rate on all the provider paths on our testbed. We find that in general, the loss rates on our testbed paths are very small. For most of the time during our measurement, our testbed paths do not experience any packet losses. In addition, the anecdote that the paths via Abilene network usually experience less packet losses than commercial ISPs does not always hold, as we observe that some Internet paths that traverse the commercial providers may actually experience much less loss than those that traverse the Abilene paths.

UMass	Loss rate (%)	UMN	Loss rate (%)	UPenn	Loss rate (%)
ma-mn-ab	0.0685	mn-ma-ab	0.0117	pa-ma-ab	0.0220
ma-mn-co	0.0449	mn-ma-co	0.0188	pa-ma-co	0.0738
ma-pa-ab	0.0319	mn-pa-ab	0.0058	pa-mn-ab	0.0529
ma-pa-co	0.0333	mn-pa-co	0.0156	pa-mn-co	0.1262

Table 1. Average loss rate on all the provider paths in the testbed.

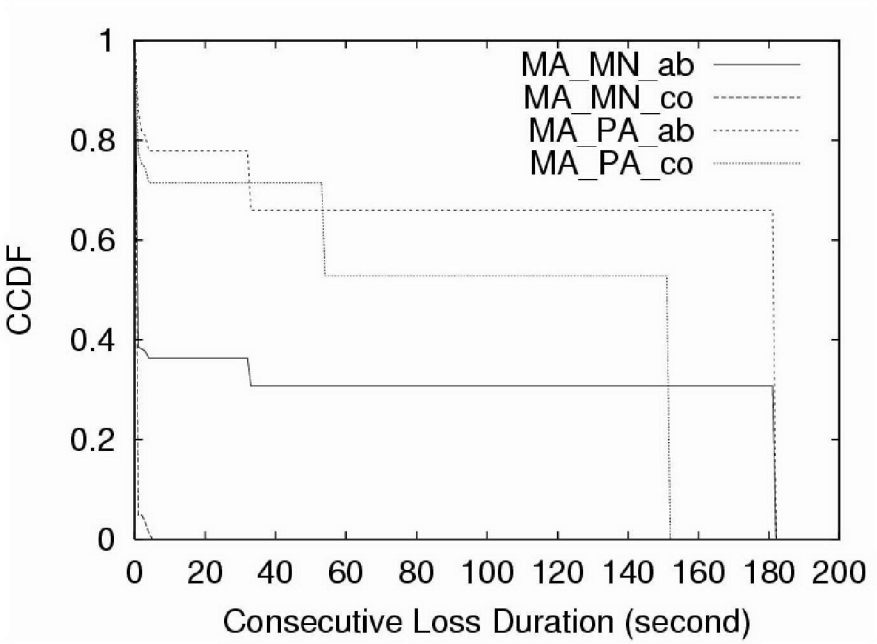


Figure 2. Complementary Cumulative Distribution of the Losses Percentage

On the other hand, although the average loss rate on those paths are, in general, small, the distribution of loss events are quite different. For most paths, the majority of the losses is short and last only one or two seconds. That is, the losses we observed on some of the Internet paths are quite sporadic, while losses on some other paths are bursty. We calculate the complementary cumulative distribution of consecutive loss events that are longer than a certain period. Figure 2 shows the loss CCDF for all the provider paths originated from UMass. We can observe that there are some loss events that last more than 180 seconds, such as the path from Umass to UMN via the Abilene provider. In addition, packets lost in those longer loss events may account for a significant portion of total packet loss. From Figure 1, we can see that on some of the paths, loss events that last longer than 30 seconds may account for more than 70% of the total packet losses observed on those paths.

Such an observation suggests that the characteristics of loss events in the Internet are quite versatile, that the single loss model or loss avoidance techniques may not be applied on all of them. This is a two-fold observation. On one hand, it means we can achieve relatively significant loss improvement by only targeting those long duration loss events; on the other hand, it means the loss prediction technique needs to be accurate. One or two mis-predicted loss events could significantly affect the performance of loss avoidance algorithms.

2.5 Correlation between loss and BGP updates

Next, we investigated the temporal correlation between the packet loss and route instabilities on our testbed. Inter-domain routing instability information may often be inferred from the BGP routing messages. We first correlate the loss events with BGP routing updates, from the BGP updates collected at two of the three sites in our testbed (Univ. MN and Univ. MA).

$$R_{xy}(\tau) = \frac{E[(x(t) - \mu_x)(y(t-\tau) - \mu_y]}{\sigma_x \sigma_y} \qquad (1)$$

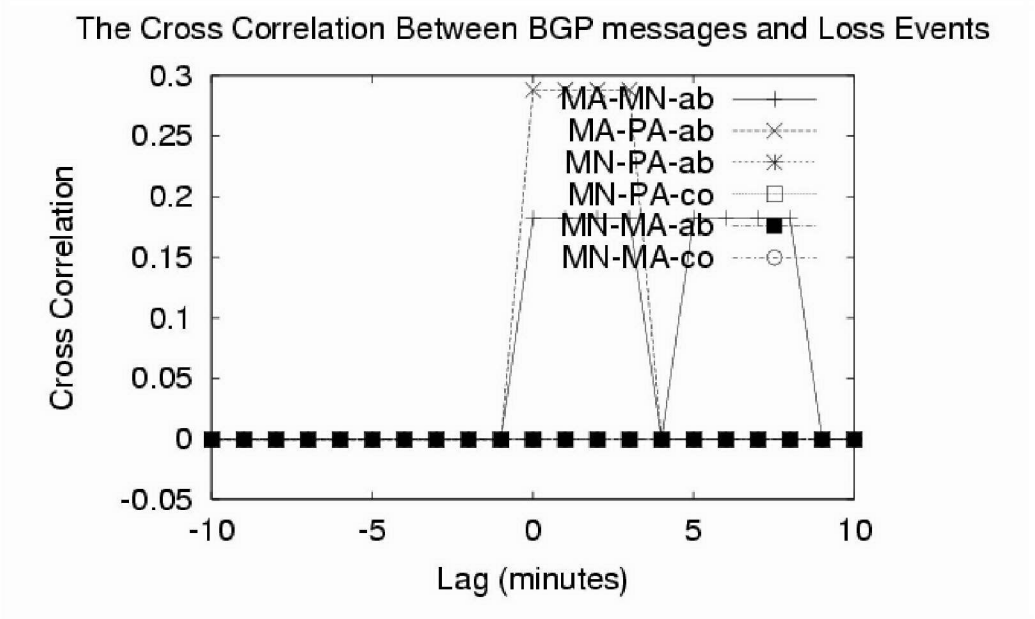


Figure 3. Cross Correlation Between Packet Losses and BGP Updates

The temporal cross-correlation between two random processes $X(t)$ and $Y(t)$ can be calculated according to Equation (1). We denote the packet loss and BGP update as two separate random processes $L(t)$ and $B(t)$. For the packet loss process, $L(t) = 1$ if the packet loss rate during an interval is greater than threshold, and $L(t) = 0$ otherwise. For the BGP process, $B(t) = 1$ if there is a relevant BGP message during observing interval t, and $B(t) = 0$ otherwise. Using Equation (1), we calculate the cross correlation between $L(t)$ and $B(t)$ and the results are shown in Figure 3.

We observe that for some of the paths, it is clear that there is much stronger correlation between the two random processes when time lag τ becomes small. In general, the cross correlations reach a maximum when the time lag $\tau \geq 0$, which means the BGP messages appear either at the same interval as the packet loss events or trail the packet loss by several minutes. Such an observation is sensible in that when routing becomes unstable on an Internet path, it usually takes some time for the routing updates to propagate back to the source. In particular, the route dampening mechanism built in the BGP protocol may further delay the propagation of BGP messages. Our observation also confirms the slow convergence phenomenon observed in [6]. On the other hand, some other paths do not exhibit varied correlation with the change of time lag τ. The correlation between BGP update and loss events is consistently close to zero.

Such a discrepancy among different paths may be explained by the loss characteristics of each path. The paths that exhibit increased correlation experience persistent packet losses that last longer than dozens of seconds, such as the path from UMass to UMN via the Abilene provider. On the other hand, those paths that contain only sporadic losses

(loss events do not last longer than 30 seconds) usually exhibit consistent near-zero correlation.

The fact that BGP messages usually follow packet loss events has important implications. Although it is possible to use BGP messages as the indication of the occurrence of bursty loss events, it is not feasible for those BGP messages to be used as the criteria to perform reactive routing in order to avoid packet loss. As the convergence of the BGP protocol is relatively slow, the loss events are usually over by the time BGP updates propagate back to the source. An example of the temporal sequence between the BGP messages and a bursty loss event is illustrated in Figure 4.

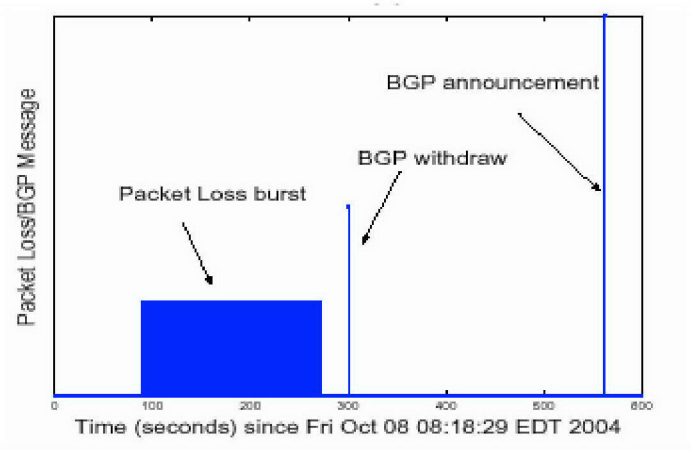


Figure 4. Temporal illustration of BGP and loss events observed by the source node (On provider path from Umass to UMN via Abilene)

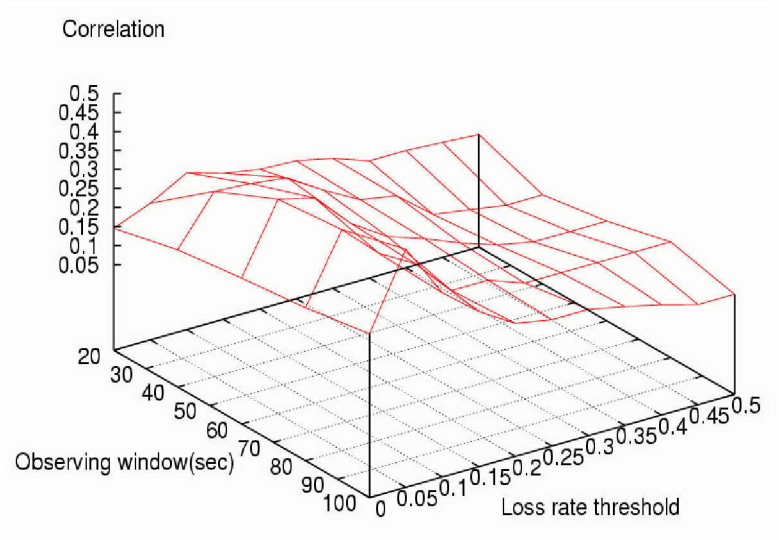
2.6 Identifying IP level path change

BGP messages summarize some of the routing instability in the Internet, but there are also other route change events that may not necessarily trigger BGP updates. In addition, BGP messages usually lag behind packet loss or failure events and have limited use for reactive routing scheme to avoid loss performance. It therefore makes sense to look at the IP level path change directly to detect route instability on an end-to-end path. In this section, we describe the heuristic that detects IP level path change from the traceroute measurements we performed on all the provider paths on our testbed.

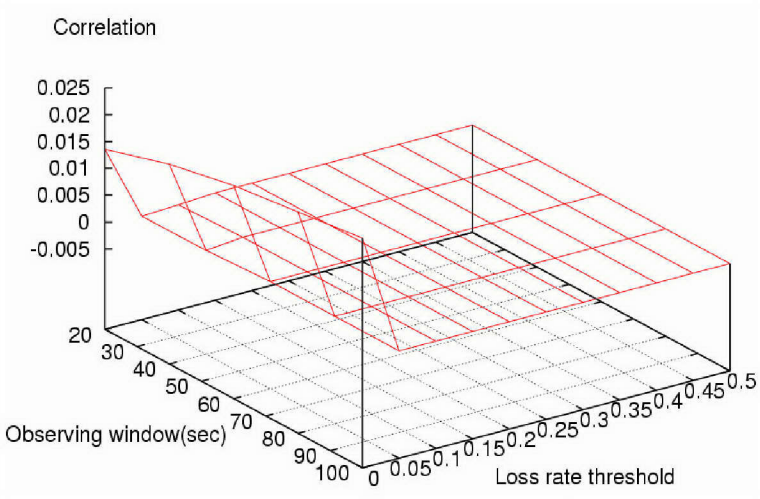
We distinguish the routing instability by labeling the end-to-end path as being in to one of two states, either "stable" or "transit". A path is considered entering the "transit" state, if a "path change" event is detected --- that is, an IP level path recorded by traceroute *i* is different from the path recorded by the previous traceroute probe *i-1*. An end-to-end path is considered entering the "stable" state, if the same IP level path is used for at least *t* seconds. In our analysis, we choose *t=60* seconds. When the end-to-end path is experiencing heavy losses, the traceroute probes may also be lost, resulting in traceroute records of incomplete end-to-end paths. In our analysis, we treat those incomplete end-to-end path records as path changes as well.

In order to make sensible identifications about routing changes, we need to filter noises in the traces that should not be identified as routing instability from our analysis. The first type of noise is due to an incomplete traceroute record. In most cases, when the Time-To-Live (TTL) value of a received packet reaches 0, routers reply to the source of the packet with a TIME EXCEEDED ICMP message. However, it is also quite common for some routers to occasionally ignore traceroute probes by not responding with ICMP messages, probably due to rate limiting reasons or a specific router implementation. In such traceroute records, the hop that corresponds to such a router will be recorded as "*" after the traceroute probe times out. We compare such an incomplete IP level path with the complete path that is most similar to it. If the hops immediately before and after the missing hop in the incomplete path are the same as those in the complete path, the missing hop must not be caused by routing change. We treat such an incomplete path the same as the complete paths in our analysis. On the other hand, if an incomplete path contains two or more consecutive "*"s, we could not reliably tell whether such missing hops are caused by routing instability. In our analysis, we treat those incomplete paths as path changes.

The second type of noise is caused by network configuration. Some ISPs adopt load-balancing techniques to deliver traffic in their network in order to improve performance or optimize the use of their network resources. Internet traffic traversing those networks may use two or three most commonly used IP level paths. Those records are traces of special network operational practices - as opposed to the routing instability that may cause performance disturbance (in which we are really interested). We manually identify the paths that traverse the network by employing such techniques and do not identify them as routing instabilities.



(a) UPenn-UMN via Commercial Provider



(b) UMN-UPenn via Abilene Provider

Figure 5. Correlation between loss and path change events with various window size w and loss rate threshold λ

2.7 Correlation between loss and IP level path changes

Having identified the path state using traceroute measurements, we hope to understand how Internet route instability affects the packet losses on the Internet paths. To this end, we start by looking at the cross correlation between the IP level path change and the end-to-end packet losses.

Using Equation (1), our analysis shows that the IP level path change reaches the highest correlation when the time gap $\tau=0$. Such observation is sensible in that both the information about IP level path change and packet loss is obtained from the data plane. Therefore, instead of temporal cross correlation, we calculate the cross correlation of the packet loss random processes $L(t)$ and the IP level path change random process $P(t)$ with various loss rate threshold λ and observation window size w.

We use the loss rate threshold λ to decide whether the Internet path is "lossy", or "not lossy". The loss random process $L(t) = 1$ if the loss rate during the observrf interval $r(t) > \lambda$, or $L(t) = 0$ otherwise. Similarly, the path change process $P(t) = 1$ if there is a positive detection of route instability during the observing window and $P(t) = 0$ otherwise. We also hope to choose the size of the observing interval w to understand how far away in time that a path change and a loss event should considered to be relevant. Therefore, we calculate the cross correlation between $L(t)$ and $P(t)$ with various observing window size w and loss rate threshold λ, and the results are shown in Figure 5.

The plot (a) shows a path with very bursty loss (from UPenn to UMN via commercial provider), and plot (b) shows a path with very sporadic loss (from UMN-UPenn via

Abilene provider). From the observations we made on all the paths in our testbed, we find that the correlations display unique characteristics for each path as well as some common traits. Our observations are summarized as follows:

- When comparing different paths, the absolute values of the correlation are quite different. For paths that contain mostly long-duration loss events, the correlation is much higher than those paths that contain mostly sporadic losses.
- When comparing the correlation on individual paths, we note the following characteristics. If the paths contain only sporadic loss events, the correlations between path change and loss events are consistently small.
- If the path contain sporadic as well as bursty losses, when the loss rate threshold λ is small, the correlation between the loss events and the routing instability is low. As the loss rate threshold λ increases, the correlation between path change and packet loss events increases accordingly. However, as the loss rate continues to increase, the correlation will eventually decrease. Such an observation is consistent across all the paths that contain bursty loss events that we studied.

The changes of correlations between path change and packet loss events on those paths are because of the following reason. When the loss rate threshold λ is small, both the bursty loss and sporadic loss events are going to be identified in the loss process. The smaller correlation indicates the path change events are not correlated very well with those loss events. As the loss rate threshold λ increases, only more bursty loss events are going to be identified in the loss process. The higher correlation suggests that the path change events are better correlated with those bursty losses. On the Internet paths we studied, there are only a few long lasting loss events with high loss rates. As the loss rate threshold λ continues to increase, not all of the path change events we identified necessarily correspond to those loss events. Therefore, the correlation between the path change and packet loss will decrease. Such observations could be seen as additional supporting evidence that long lasting loss events are more often correlated with routing instability.

3. Route-Instability Based Path Switching

Having had the evidence that long duration loss events are indeed correlated with the IP level path change, we investigate the feasibility of using reactive routing techniques to avoid those packet losses. In particular, we focus our attention on the effectiveness of performing end-to-end path switching to exploit the benefit of path diversity provided by multiple providers. To this end, we designed several algorithms to improve end-to-end loss rate as well as reducing loss burstness. In this report, we present two of those path-switching algorithms.

3.1 Related Work

In recent years, path diversity in the Internet has attracted much attention in the research community as well as being put into practice by industry. In [19 20 21], the authors investigated the characteristics of path diversity and its implications on particular Internet applications. [16 17 18] exploit Internet path diversity to provider services with performance guarantees provided to businesses as well as end users. In the MIT RON project [14], overlay routing is used to find alternative paths in the Internet, in order to improve the resilience of the network and improve availability. RON nodes examine the condition of the Internet between themselves and the other nodes, and, based upon the network condition, decide if they should send packets directly to other nodes, or send them indirectly via other RON nodes. However, [14] does not examine and respond to the effects of routing dynamics on packet loss. End-to-end path switching is another reactive routing technique that exploits the path diversity provided by multiple networks paths. In [4], the authors showed that, for a multi-homing network, paths to the Internet provided by multiple providers do not all exhibit performance degradation at the same time. Therefore end hosts are able to select the provider path route traffic with the best performance, in order to avoid the performance degradation such as packet losses on the original path.

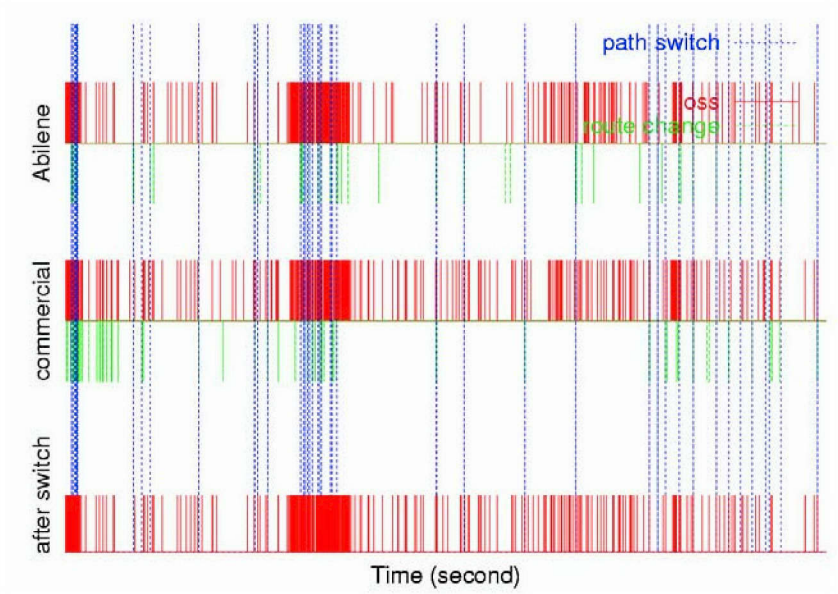
Our work uses end-to-end path switching to improve the loss performance on the Internet paths. Instead of performing path switching based on past loss observation, the intuition of our strategy is based on the observation on the correlations between the routing instability and packet loss. Our end hosts monitor the routing instability by constantly probing the IP level Internet paths. The end host routes packets via an alternative path if it detects that the state of the Internet path is "instable", thus to avoid the potential packet loss that may be caused by a routing change.

3.2 Path switching algorithms

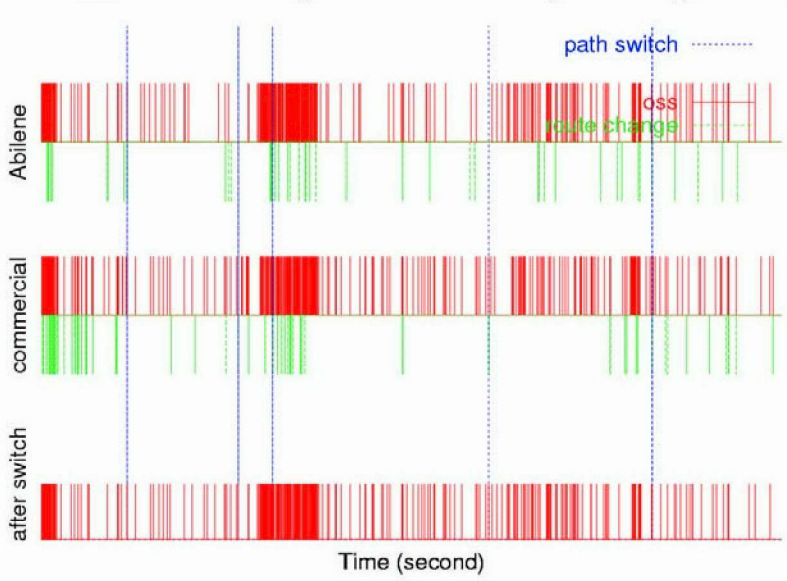
Using the traces collected from our active measurement experiment on the wide-area testbed, we emulate the process of path switching. Our focus is on the effectiveness of various path-switching algorithms in terms of improving end-to-end loss performances.

The first path-switching algorithm is based purely on the detection of routing instability from the traceroute record. The path switching strategy monitors path states on all the

parallel paths between a source and destination node, and makes a path-switching decision every *w* seconds. If the current path ever enters the "instable" state during the past observing window of length *w*, the algorithm deems that the path has entered a "bad" state and assumes the current path may potentially incur packet loss. The algorithm will switch to an alternative path, if there is at least one alternative path that has not entered the "bad" state during the past *w* seconds. If there is more than one candidate path whose path states are predicted to be good, we randomly choose one alternative path to switch to. If all the alternative paths are predicted to be "bad", the algorithm will stick with the current path.



(a) Path switching based on IP level path change



(b) Path switching based on both IP level path change and packet loss

Figure 6. The packet loss on paths before and after path switching (provider paths from UPenn to UMN)

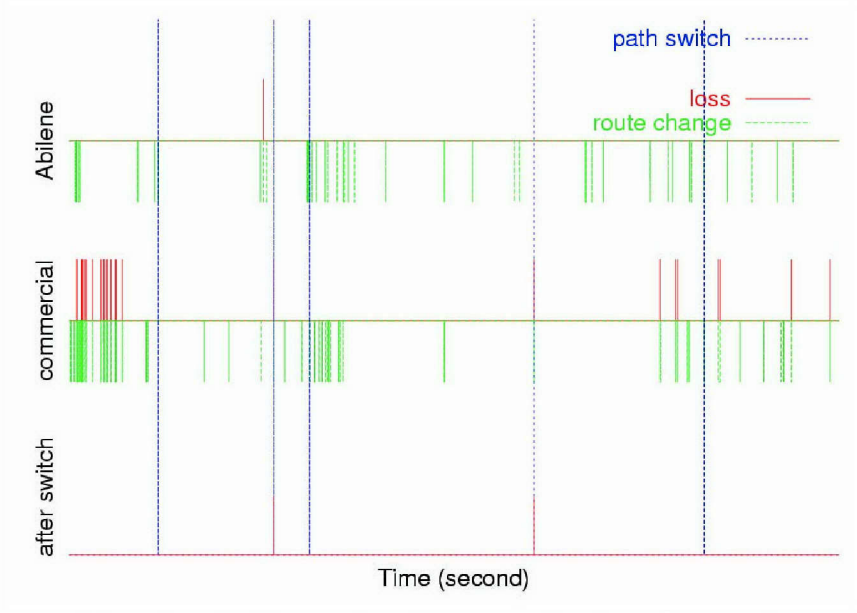
Our second path-switching algorithm uses both packet loss and IP level path change information. In order to avoid unnecessary path switching, the path-switching algorithm decides that the path has entered a "bad" state only if the current path enters the "uinstable" state and there are at least two packet losses during the past observing window of length *w*. Similar to the first switching algorithm, the second switching strategy monitors path states on all the parallel paths between a source and destination node, and makes a path-switching decision every *w* seconds. The algorithm will switch to an alternative path if there is at least one alternative path that has not entered the "bad" state during the past *w* seconds. If there is more than one candidate path whose path states are predicted to be good, we randomly choose one alternative path to switch to. If all the alternative paths are predicted to be "bad", the algorithm will stick with the current path.

The effects of the two path switching algorithms are shown in Figure 6. In Figure 6 (a), the above two plots show the packet loss (identified with red solid bars) and path change events (identified with green dash bars) over time on two provider paths between UPenn and UMN before path switching. The plot on the bottom shows the packet losses on the paths after path switching. The long blue dot bars identify the time at which path-switchings occur. Figure 6 (b) shows a similar plot for the second path-switching algorithm. From Figure 6, we can see that path-switching algorithms based on both loss and path change information performs much fewer switches than the algorithm that is based on IP level path change information alone.

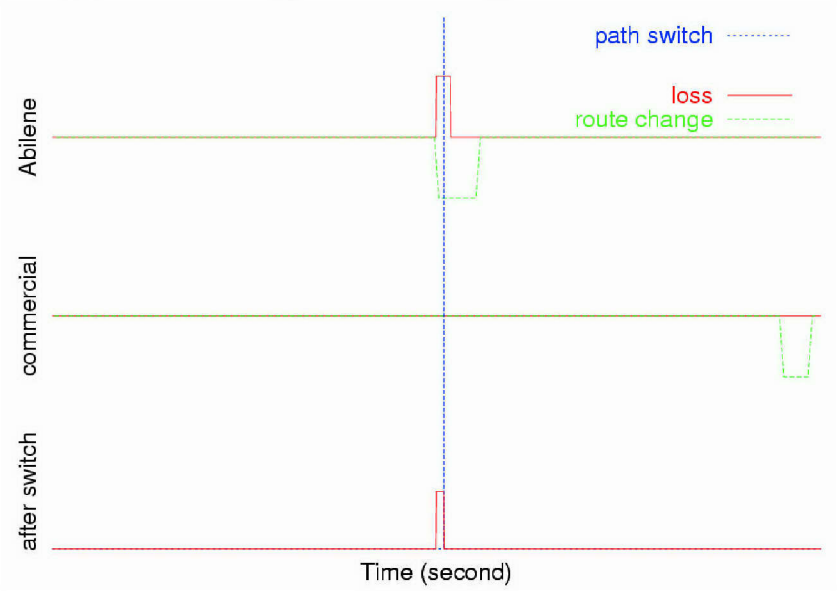
MA-MN	Loss rate (%)	1st (%)	2nd (%)	MA-PA	Loss rate (%)	1st (%)	2nd (%)
ab	0.0685	0.0513 (42)	0.0459 (6)	ab	0.0319	0.0153 (16)	0.0204 (4)
co	0.0449			co	0.0333		
MN-MA	Loss rate (%)	1st (%)	2nd (%)	MN-PA	Loss rate (%)	1st (%)	2nd (%)
ab	0.0117	0.0119 (27)	0.0131 (2)	ab	0.0058	0.0087 (37)	0.0073 (1)
co	0.0188			co	0.0161		
PA-MA	Loss rate (%)	1st (%)	2nd (%)	PA-MN	Loss rate (%)	1st (%)	2nd (%)
ab	0.0220	0.0221 (8)	0.0209 (3)	ab	0.0529	0.0596 (47)	0.0542 (6)
co	0.0738			co	0.1262		

Table 2. The average loss rate before and after path switching on all the provider paths

The average loss rates achieved by each switching algorithm are listed in Table 2. The 1st column records the loss rate after the 1st path-switching algorithm, and the 2nd column records the loss rate after the 2nd path-switching algorithm. The number in the parenthesis indicts the number of switches performed by each algorithm. It is noticeable that the 2nd switch algorithm achieves comparable results by performing much fewer switches.



(a) Path switching on loss burst equal at least 2 seconds



(b) Path switching on a single bursty loss event

Figure 7. Illustration of path switching in reducing loss burstiness on provider path from UPenn to UMN

(Switch based on both path change and loss information)

We then investigate the performance of our path switching algorithms in terms of reducing loss burstiness. We look into only those loss bursts that are longer or equal to certain durations. Figure 7 (a) shows that the path switching algorithm based on both path change and loss information can respond to all the bursty losses that last longer than 2 seconds and switch to a better path. In addition, the path-switching algorithm can

reduce the duration of a single loss burst, as shown in Figure 7 (b). An example of the complementary cumulative distribution function of consecutive losses is shown in Figure 8 (Provider paths between UPenn and UMN).

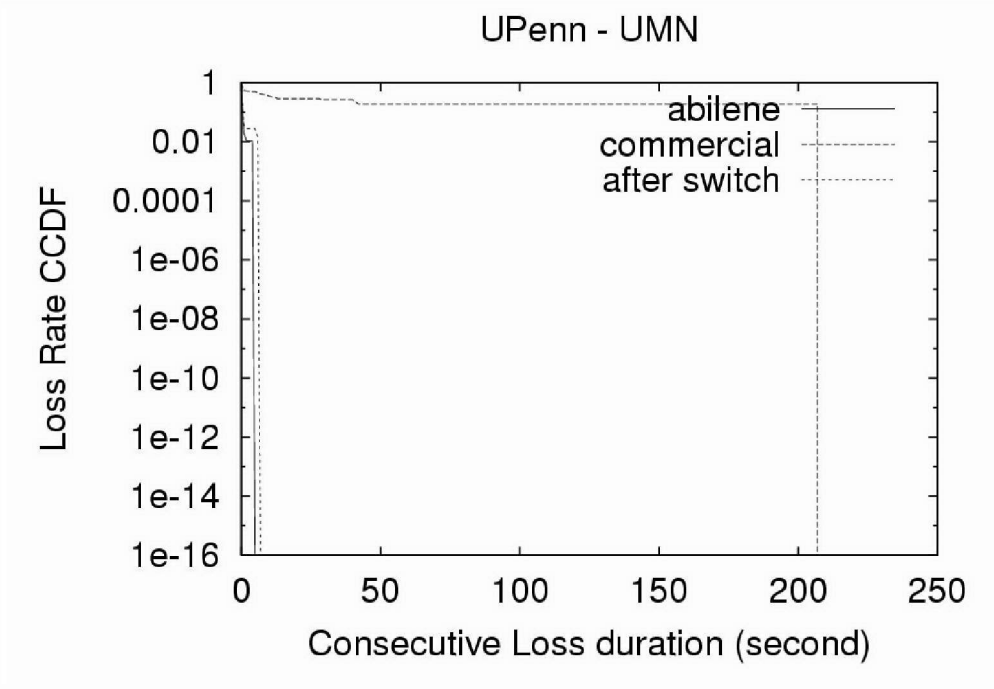


Figure 8. Loss Rate Complementary Cumulative Distribution For all Paths After Switching Based on Both Route Instability and Loss (UPenn-UMN)

Our results show that the performance of the path-switching algorithm is affected by the loss characteristics of the candidate paths. Sporadic losses are very hard to predict and avoid. Both the IP level path change and the packet loss information are obtained via the data plane along the same path as the probing packets. When the path-switching algorithm detects the IP level path change, and tries to react to the path change, the sporadic loss events are already over. The prediction based path switching will not achieve noticeable gains even though the algorithm performs switches along those paths many times. Sometimes it even produces traces that are a little burstier than one of the original traces. In such cases, however, it worth noting that the performance degrading is tolerable, as the longest bursty loss length is still very small. On the other hand, the path-switching algorithm performs much better for bursty losses. The route instability based path switching algorithm consistently achieves satisfactory performance gain when both the original paths contain relatively long duration loss events (longer than 10 seconds). And the longest loss burst after path switching never exceed more than 40 seconds for all the traces we analyzed. We also note that for this path-switching algorithm, it performs many switches that do not contributed towards helping reduce average loss rate or loss burstness.

Of all the path-switching algorithms we proposed and investigated, the switching strategy based on both routing instability and packet loss seems to achieve the best results. In essence, this algorithm performs path switching from two perspectives. From a loss-

based path switching perspective, it tries to reduce the number of path switches by ignoring the low loss rate events that occurred on the path, but compensates with the detection of IP level path change. From a path change based switching perspective, it tries to respond to the route instability events that actually cause packet loss, therefore increasing the accuracy of our route instability based prediction by taking into consideration of those high loss rate events.

4. The Routing Loops in the Internet

We initiated this study to investigate the problem whether Internet topology as well as routing behaviors may contribute to potential problems of the Internet. Some abnormal routing behaviors caused by current routing practices, such as the routing loops may cause packets of the Internet traffic to be delivered to un-desired destinations, and consume unnecessary network resources. Such a vulnerability, although usually short-duration and may not be obvious or even noticeable in normal operation, may exacerbate under the stress of malicious computer attacks, such as computer virus outbreak or Distributed Deny of Service (DDoS) attack.

We focus on the Internet worm outbreaks scenario and study the impact of traffic overhead on the network caused by worm propagation. A computer worm propagates itself by scanning IP addresses in the Internet, detecting those machines that are vulnerable, and trying to duplicate and install itself onto the target. Once infected, the victim machine is going to repeat the process and try to infect other vulnerable machines in the Internet. Such a process carried out by tens of thousands of machines may easily cause network congestion or overwhelm the processing power of routers.

In order to model the network behavior under stress, we investigate the current Internet architecture as well as the existing anomalous Internet routing behavior.

4.1 Related Work

The topology of the Internet has been one research topic that has caught much attention [4] [5] [10]. In particular, there have been several studies that focus on Internet routing loops [7] [8]. Those studies, however, mostly focus on understanding the architecture or routing behavior in normal operational situations, while our study focuses on the impact of those behaviors in abnormal operational situations and how they might affect or disrupt the normal network operations. In [8], the authors made an observation on routing loops that exist in the Internet by analyzing packet trace collected from a major ISP's backbone network. Our observation, however, was made based on active end-to-end measurement and therefore may reveal a routing loop that is beyond the scope of a single network's management.

In order to understand the impact of the abnormal routing behaviors, we first make observations on the basic characteristics of the Internet topology and routing behaviors.

4.2 AS level Internet Characteristics

The simplest worm scanning method is called the random scanning. Namely, the worm program randomly scans IP addresses selected from the 2^{32} IP address space with equal probability. In such a scenario, the number of IP addresses owned by a network is proportional to its possibility of being scanned by a computer worm, and therefore the amount of scanning traffic it may receive. The effect of such scanning traffic is also related to the number of connections that the traffic uses to reach the network.

Previous studies have shown that the connectivity of the networks, a.k.a. AS degrees, conforms to the power law distribution. We verify such a result with our own analysis of the global routing table. Such a distribution indicates the majority of the networks in the Internet have only a few number of connections with other networks, and the networks with very high AS degrees are rare.

We also looked into the distribution of IP address spaces owned by individual networks. We found that, while networks with a large number of IP addresses are rare, networks with small IP address spaces are also much less compared with the case of AS degree. This may contribute to the fact that networks with very few addresses usually do not qualify to be an individual autonomous system.

We observed that the networks with large IP address spaces do not necessarily have high AS degrees and vice versa. The effect of worm scanning traffic on an inter-AS link would be affected by the address space covered by prefixes using that link. In principle, those networks with large IP address space but with few connections with other networks are more susceptible to the worm scanning traffic. A detailed investigation suggests that such networks are often state or regional networks, or networks in other continents. We investigated our conjecture that such networks would experience the most significant impact in a malicious attack scenario.

4.3 Measurement

We use a measurement approach to study the behaviors of routing loops in the Internet. We performed end-to-end traceroute to different destinations in the Internet from our server located in the University of Massachusetts.

The destinations of our probing are chosen from the routable address blocks (RAB) covered in the global BGP routing tables, as well as those nonroutable address blocks (NRAB) that are not covered. For each set of aggregated prefixes, we probe two IP addresses, one at the beginning of the address block, the other at the middle of the address block. Offline analyses are used to categorize routing loops recorded in the trace data.

The measurement data were collected during a period of 3 weeks, and the preliminary results we obtained are as follows:

- Prevalence of Routing Loops

  We found that 7.3% of the RAB traces were identified with routing loops of size 2 and more. Another 8.1% of the traces were identified with the repetition of a single IP address. Together, they accounted for 15.4% of the total measurements performed. Inter-domain routing protocols such as BGP, intra-domain routing protocols (IGP) such as IS-IS or OSPF, static route configuration, or a combination of these may all contribute to those routing loops. For non-routable addresses, the number of traces containing possible routing loops is almost negligible.

- Persistency of Routing Loops

  Some of the routing loops that we observed persist over a long period of time and may cause permanent routing problems.

  Our analysis showed that more than 80% of the routing loops identified in our first measurement lasted more than one week. Moreover, the percentage is even higher if we only compare the loops whose sizes are 2 and above. This result shows that routing loops in the Internet last longer than previous believed.

- Span of persistent Loops

  Our study paid special attention on those persistent routing loops whose sizes are 2 and above. Of the total of 2634 such loops, 91% are confined within a single AS, while more than 8.6% of the persistent loops span across multiple ASs. The router level sizes of the loops confined in a single AS range from 2 to 6, while the size of the loops span across multiple ASs may have router hops as large as 18.

- Location of persistent Loops

  Our study also shows that for persistent loops confined within a single autonomous system (AS), over 70% of the time they fall in the same AS as the destination network. For persistent loops that span over multiple ASs, over 40% of the time the loops start in the same AS as the destination. A closer look reveals that for those persistent loops not reaching the destination ASs, 68% reach an AS that is the direct neighbor of the destination autonomous systems.

5. Summary

In this study, we performed end-to-end measurements on a wide-area network testbed to understand the loss characteristics on the Internet paths. We showed that bursty losses accounts for significant portion of the total loss observed by the end-hosts. Those bursty losses can often be correlated with Internet routing instability and we verified the results using traceroute and BGP routing information. Having made those observations, we showed that it is possible to exploit the benefit of path diversity to improve the end-to-end loss performance. We proposed several end-to-end path switching algorithms to reduce the loss burstness and overall loss rate, based on the Internet routing instability information.

Our results show that routing instability based path switching algorithm with simple prediction strategy can produce satisfactory results in most cases. Although our observation was made on a small sized testbed, we deem this work as the initial step to a effective and practical dynamic path switching system.

We also initiated the investigation on the impact of malicious traffic on the Internet. We hope to understand how the current Internet is going to be affected by malicious traffic and how the existing routing anomalies such as routing loops would exacerbate under stress. Our current findings are useful and serve as an initial step to design a realistic model to discover and study those problems of the Internet.

From our measurements, we found that the majority of the routing loops are limited within a single AS, and the impact on traffic overhead caused by routing loops are likely to be confined in individual networks. Inter-domain links would only be affected by traffic caught in inter-domain routing loops, which are rare events. However, the fact that most of the routing loops happen near the destination network indicates that all the inter-mediate networks would unnecessarily carry traffic along the way and such overhead may be potentially reduced. Our analysis also showed that the amount of worm scanning traffic imposed on an inter-domain link would be proportional to the addresses covered by the prefixes using that link, and that those large networks behind only a few links connecting to the rest of the Internet are more likely to experience performance degradation such as congestions. State and regional networks in the United States or national backbones in other continents are more likely to have such a network topology, and may be more susceptible to Internet worm attacks.

We used an active measurement approach to study both problems. For the end-to-end loss performance improvement, although our observation was made on a small sized testbed, we deem this work as the initial step to an effective and practical dynamic path switching system. Our next step is to investigate the feasibility of employing more advanced reactive routing mechanism such as overlay routing. For future work we will expand the measurements to the global Internet and analyze the Internet routing loop characteristics to understand the cause of such loops.

6. References:

[1] M. Yanjnik, J.Kurose, and D. Towsley. Packet loss correlation in the Mbone multicast network: experimental measurements and markov chain models. *Technical Report, Umass CMPSCI TR#95-115*

[2] M. Yanjnik, S. Moon, J. Kurose, and D. Towsley. Measurement and modeling of the temporal dependence in packet loss. In *Proc. of IEEE INFOCOM,* March 1999

[3] J. Bolot. End-to-end packet delay and loss behavior in the Internet. In *Proc. Of ACM SIGCOMM*, September 1993

[4] Shu Tao, Kuai Xu, Ying Xu, Teng Fei, Lixin Gao, Roch Guerin, Jim Kurose, Don Towsley, Zhi-Li Zhang. Exploring the performance benefits of end-to-end path switching. in *Proc. of IEEE ICNP* 2004.

[5] Nick Feamster, David G. Andersen, Hari Balakrishnan, and M. Frans Kaashoek. *Measuring the effects of Internet path faults on reactive routing* ACM SIGMETRICS 2003.

[6] Labovitz, C., Ahuja, A. Bose, A and Jahanian, F. Delayed Internet routing convergence. *IEEE/ACM Transactions on Networking,* 2001

[7] L. Subramanian, S. Agarwal, J. Rexford and R.H. Katz Characterizing the Internet hierarchy from multiple vantage points *IEEE INFOCOM,* 2002

[8] Aditya Akella, Srinivasan Seshan and Anees Shaikh An Empirical Evaluation of Wide-Area Internet Bottlenecks *Internet Measurement Conference (IMC)* 2003, Miami, Florida

[9] David Meyer University of Oregon Route Views Project http://www.routeviews.org/

[10] V. Paxson End-to-End Routing Behavior in the Internet *IEEE/ACM Trans. Networking,* 1997

[11] Ashwin Sridharan, Sue Moon and Christophe Diot On the Causes of Routing Loops *ACM IMC* 2003

[12] L. Gao On inferring Autonomous System Relationships in the Internet *IEEE/ACM Trans. Networking* 2001

[13] Michalis Faloutsos, Petros Faloutsos, Christos Faloutsos On Power-Law Relationships of the Internet Topology *ACM SIGCOMM* 1999

[14] MIT RON Project. http://nms.lcs.mit.edu/ron/

[15] A. Feldmann, O. Maennel, Z. M. Mao, A. Berger, B. Maggs, Locating Internet Routing Instabilities, In *Proceedings of ACM SIGCOMM*, 2004.

[16] Internap. http://www.internap.com, November, 2004

[17] Routescience. http://www.routescience.com, November, 2004

[18] Skype. http://www.skype.com, November, 2004

[19] A. Akella, B.Maggs, S.Seshan, A. Shanikh, and R. Sitaraman. A measurement based analysis of Multihoming. in Proc. Of ACM SIGCOMM, August, 2003.

[20] D. G. Andersen, A. C. Snoeren, and H. Balakrishnan. Best-path vs. multi-path overlay routing. In *Proc. of Internet Measurement Conference*, October, 2003.

[21] R. Teixeira, K. Marzullo, S. Savage, and G. M. Voelker. In search of path diversity in ISP networks. In *Proc. of Internet Measurement Conference*, October, 2003.